DOCUMENT RESUME

ED 420 684 TM 028 359

AUTHOR Fenton, Ray; Straugh, Tom; Stofflet, Fred; Shaffer, Frank
TITLE EXCELS!--Evaluating Integrated Language Arts and Social Studies. The Jamestown Integrated Performance Assessment Unit.
INSTITUTION Anchorage School District, AK.
PUB DATE 1998-04-00
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998).
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Colonial History (United States); *Curriculum Development; *Educational Assessment; *Evaluation Methods; Grade 4; Integrated Activities; Intermediate Grades; Language Arts; *Performance Based Assessment; Social Studies; Standardized Tests; Student Evaluation; *Test Construction
IDENTIFIERS *Anchorage School District AK; Virginia (Jamestown)

ABSTRACT

The EXCELS! Project of the Anchorage School District (Alaska) was initiated in 1994 through a grant from the U.S. Department of Education to demonstrate the development of a kindergarten through sixth grade instructional program reflecting state and national standards in History, Civics, Geography, and English. More than 1,200 Anchorage teachers and librarians participated in the EXCELS! training and curriculum development activities. A multifacted assessment was later undertaken under the direction of the Anchorage School District Assessment and Evaluation Department. This paper reports on a particular performance-based assessment for grade 4, the Jamestown Assessment Unit. This unique integrated unit focused on U.S. colonial history and the settlement of Jamestown (Virginia). Over the 2 years of development, more than 70 people contributed to the final development of the Jamestown Assessment Unit. Fifteen teachers and more than 400 students field tested elements of the assessment and scoring procedures. All fourth grade teachers know that they may be called on to offer this unit as part of the Anchorage Assessment Program. In spring 1997, 412 students participated in the Jamestown Assessment Unit. Statistical analyses of the data from 368 students show that student work generated as part of instruction may be used for a standardized assessment of student and program performance relative to local, state, and national standards. The Jamestown Unit is an operational definition of an authentic assessment, scorable by both trained raters and teachers. (Contains 22 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 420 684

ANCHORAGE SCHOOL DISTRICT

EXCELS! -- Evaluating Integrated Language Arts and Social Studies

THE JAMESTOWN INTEGRATED PERFORMANCE ASSESSMENT UNIT

APRIL 1998

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY HAS BEEN GRANTED BY
Ray Fenton
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
2A

ASSESSMENT AND EVALUATION DEPARTMENT

Ray Fenton
Tom Straugh
Fred Stofflet
Frank Shaffer

TM028359

EXCELS! -- Evaluating Integrated Language Arts and Social Studies
The Jamestown Integrated Performance Assessment Unit

Ray Fenton, Tom Straugh, Fred Stofflet and Frank Shaffer[1]
Anchorage School District

Introduction

The EXCELS! project was initiated in 1994 through a grant from the United States Department of Education to demonstrate the development of a K-6 instructional program reflecting state and national standards in Civics, Geography, History, and English. The expectation was that student knowledge and skill would be enhanced through the infusion of reading, writing, speaking, and listening activities into social studies instruction.

More than 1,200 Anchorage teachers and librarians participated in EXCELS! training and curriculum development activities between 1995 and 1998. A core team of teachers was introduced to standards based instruction and curriculum development through innovative multi-discipline summer institutes. These teachers created school-based teams to diffuse what they learned through their in-service training and to support curriculum and lesson development in their individual schools.

Summer institute participants and selected teachers from participating schools were exposed to national leaders in the development of standards and curriculum. Educational leaders such as Christopher Salter, Cathy Salter, Miriam Chaplin, Elaine Reed, Beth Farnbach, Jack Hoar, Kathy Short, and Paul Gagnon contributed inspirational insights and connections with current national efforts to improve education. Local leaders such as Doug Phillips, Rebecca Sipe, Mary Bristol, and Barbara Bernard provided the project with day-to-day leadership and supported the development of curriculum and instructional materials. Ray Fenton directed the local evaluation effort with the assistance and support of the external program evaluator, Dr. Sandra Stotsky.

A multi-faceted assessment effort was undertaken under the direction of the Anchorage School District Assessment and Evaluation Department. The assessment focused on the impact of EXCELS! program on curriculum, knowledge of teachers, instruction provided to students, and student achievement. Teachers provided information on their training, experience, and changes in instruction. University of Alaska based researchers examined the impact on the Anchorage School District curriculum, school level

---

[1] The Jamestown Assessment Unit could not have been produced without the able assistance of Judy Ware, Joe Lisenby and Shana McPheters. This manuscript benefited from the review and comments of Dr. Steve Garrison of the Anchorage School District.

programs, and classrooms through a combination of focused interviews, site visits, and classroom observations. The Assessment and Evaluation Department tracked norm referenced test scores in reading, language arts, and social studies as well as developing a unique integrated assessment unit directly related to the goals and expectations for student gains in both language arts and social studies.

The overall evaluation results were good. Teachers gave positive reports of their experience and indicated that they were doing more to incorporate standards based instruction into their classrooms. University based observers and the program's external evaluator saw positive systemic changes when they visited schools, interviewed principals, and observed instruction. Norm referenced test results in reading, total language arts, and social studies showed gains across the Anchorage School District with the greatest gains in those schools which were early adopters of the EXCELS! project.

It was realized early in the project that an objective measure of student performance directly tied to standards and instruction would be needed to make a definitive statement about student learning. It was also hoped that a well made authentic assessment system could model the integration of instruction expected in the EXCELS! project and have a positive impact on classroom practice (Falk and Darling-Hammond, 1993). This paper reports on the development of the Jamestown Assessment Unit and the performance of students on this unique integrated grade four performance assessment.

The Birth of the Jamestown Assessment Unit

The difficulty of assessing student performance related to standards based expectations in reading, writing, speaking, and listening within the context of the social studies was recognized from the onset of the project. The direct assessment question from the federal grant which drove the development of the performance asked, "Did student products increasingly reflect the national standards and the content and performance skills in Civics, History, Geography, and Language Arts?"

During the summer of 1995 two groups of EXCELS! project participants met to identify the critical student skills in the language arts and social studies areas that related to national, state, and local standards. The social studies group developed a set of two dozen specific areas where students should be expected to develop skills in the social sciences. These ranged from map making and research skills through the ability to synthesize information, provide sound critical judgments, and prepare reasoned arguments based on observation and the application of theory. The language arts group adopted the Alaska State Language Arts Standards which call for A) speaking and writing well for a variety of purposes and audiences, B) being a competent and thoughtful reader, listener, and viewer, C) having the ability to identify and select from

multiple strategies to complete projects, D) thinking logically and reflectively to present and explain positions based on relevant and reliable information, and E) demonstrating the ability to understand and respect the perspectives of others (Alaska Department of Education, 1996). The student skills were formalized into a set of statements by Neil O'Leary of Dos Gatos Enterprises and presented to the EXCELS! Core Committee.

The EXCELS! Core Committee made up of Anchorage School District language arts and social studies curriculum coordinators, the EXCELS! project directors, and assessment and evaluation staff identified grade four as the target grade for an integrated assessment unit. A "skills matrix" was prepared to serve as the guide for the development of the performance assessment. The general direction given to the development team by the EXCELS! Core Committee was that the content of the assessment would be drawn from social studies and that the performances collected would be based on reading, writing, listening, and speaking.

The content of the fourth grade curriculum and curriculum materials was reviewed and a team of graduate students from the University of Alaska, Frank Shaffer and Anna Bonderinko. The specific social studies skills identified after an analysis of the K-4 curriculum were listening and note taking, map making, retrieval of information, description, analysis and synthesis of information, and the justification and explanation of conclusions based on social science theory.

The Assessment and Evaluation team then developed a number of scenarios around which an assessment unit might be developed. The areas consistent with the grade four curriculum were the early Native cultures of the Americas, the age of European exploration and explorers, contact between European and American Natives, and the early development of the Dutch, English, Spanish and Portuguese settlements in the Americas. A decision to place the assessment in the third or fourth quarter of the year further limited the content of the assessment to exploration, contact, and colonization.

The early settlement of Jamestown (1609) was selected as the content of the unit because of the availability of curriculum materials and consistency of the early settlement with curriculum goals. The Division of Historical Interpretation and Preservation at the Colonial National Historical Park provided bibliographical information and copies of interpretive materials including student and teacher materials for classroom enrichment, maps, and other images. The staff at Jamestown NPS Education Support took photographs of artifacts on display at the park and provided permission for the use of the images in the assessment. The Christian Science Monitor extended copyright permission to use text from a series of articles published on November 19, 1997 about Jamestown. The National Center for History in the Schools allowed the use of their background information for teachers

from a grade five unit on Jamestown which they have available as a demonstration of a standards based instructional unit (Pearson and Watkins, 1991). Additional information on the National Historical Park, Jamestown, and the settlement of Jamestown was gathered through the internet from the National Park Service and the Library of Congress.

Ray Fenton and the University of Alaska graduate student team visited fourth grade classrooms, observed language arts and social studies lessons, and discussed the potential characteristics of an instructional unit with teachers. Ray Fenton, Tom Straugh, and the graduate students then examined the literature on performance assessments in social studies and language arts. Judy Arter of the NWREL test center and the ERIC Clearinghouse on Assessment and Evaluation provided a number of examples of performance assessments. Some suggestions and models related to curriculum were provided by Gary Nash of the National Center for History in the Schools and the National Council on Economic Education. No assessment was found that would serve as a complete model for a standards based assessment which would directly reflect Anchorage School District, State of Alaska, and national standards.

After extended discussion of the design of the assessment unit between the EXCELS! Core Team and assessment staff, it was decided that the unit would consist of a number of lessons with imbedded assessments consistent with the social studies content following a model provided by Richard Shavelson for the development of science assessments (Shavelson, 1995). Student work would be collected in a student journal similar in form to a laboratory notebook. Assessment tasks were reviewed for their linkage to curriculum, focus, simplicity, and propriety (ASCD, 1995).

The completed assessment unit and supporting materials were prepared and submitted to the Anchorage School District for acceptance as supplemental instructional materials consistent with the approved district curriculum. Because the assessment is considered part of the regular district curriculum, teachers may use student scores on the assessment as basis for student grades.

Through two years of development more than seventy individuals contributed to the creation of the final Jamestown Assessment Unit. There were sixteen members on each of the Language Arts and Social Studies Objectives development teams that recommended the specific objectives to be assessed and the best way to measure student performance on those objectives. Five members served on the EXCELS! core team which provided oversight and final approval of the assessment. Seven individuals participated on the actual test development team. Fifteen teachers and more than four hundred students field tested elements of the assessment and scoring procedures. Fifteen teachers, assessment, curriculum, and university staff members participated on a field test review team and scoring group.

Three staff members worked as the primary production group in making Jamestown Assessment Unit Kits. Key members of the Assessment and Evaluation and EXCELS! Core Team were trained with the ASCD Assessing Student Performance Inquiry Kit (Arter, 1991).

The Jamestown Assessment Unit

Each teacher is provided with a Jamestown assessment unit kit. The kit includes the materials needed to support the assessment unit including:

- **Jamestown: The Beginning**. This is a colorful sixteen minute video on the founding and first years of colonial Jamestown told from the point of view of an early settler (Jamestown - Yorktown Historical Society, undated).
- Colonial Jamestown Integrated Assessment Unit: Teachers Guide. This forty-page publication includes background information on Jamestown, outlines of lessons, standard instructions for the data collection elements of lessons, and scoring rubrics and keys.
- Artifact kits including photographs from the Jamestown National Historical Park to be used with the small group activity.
- Jamestown Journals. Each student keeps a Journal during the Jamestown Unit which includes both instructional sections which match lessons and the selected response, short answer, and extended response items which are the basis of assessment and grading.

The final design of the Jamestown lessons is presented in the table below along with the materials provided and the standards addressed. There is no fixed length for any lesson through some of the assessment activities have specific directions which suggest time limits. Teachers are encouraged to extend and elaborate on lessons as much as they feel necessary. The actual teaching time of the lessons is from 30 minutes to 90 minutes depending on the activities. Some lessons may last more than one day. Teaching time in the pilot and first year assessment groups ranged from five to eight days. Some teachers developed additional lessons which included research on Jamestown, map skills (presented after the map skill assessment), use of the WWW for research on Jamestown and colonial America, and the more general history of Dutch, English, and Spanish settlement in the United States.

All grade four teachers know that they may be selected to offer the Jamestown Assessment Unit as part of the District-wide Anchorage Assessment Program. Specific teachers are randomly selected for the annual district assessment. Teachers are warned at least six weeks in advance that they have been selected to give the assessment during the fourth quarter of the year. Teachers who are selected are offered a half day of substitute teacher assistance to complete the paperwork related to the unit and the opportunity to participate in the

District-wide scoring of performance assessments during the following summer. Teachers are expected to grade the work of each student using the rubrics provided in the teacher's guide which accompanies the unit.

Description of the Jamestown Assessment Unit

Lesson	Teacher's Materials	Assessment
Lesson 1 Introduction to Colonial Jamestown Goals: -Introduction to Vocabulary -Introduction to location and goals of the Jamestown colony -Building interest in Jamestown -Linking with prior lessons on exploration and contact	Lecture Video	No Assessment
Lesson 2 A letter to Martha -Introduction to a child's life in early colonial America -Learning the physical location and principal occupations of Jamestown	Text material Poster/Map Discussion Guide	Jamestown Journal -List of significant details from listening assignment. -Map of Jamestown
Lesson 3 The New World News -Learning about the Virginia Company -Exploring the reasons that Europeans came to colonize North America -Examining the economics of early colonies -Examining the relations between Native Americans and colonists	Text materials - newspaper Discussion Guide Enrichment Materials	Jamestown Journal -Selected response reading comprehension - Constructed response synthesis and interpretation of reading.

Description of the Jamestown Assessment Unit (continued)

Lesson	Teacher's Materials	Assessment
Lesson 4 American Anthropologist -Introduction to Anthropology -Classification, examination, and reasoning about artifacts -Activities of colonial Jamestown.	Text materials Group Activity Kits Student response sheets Discussion Guide Poster Constructed Response Assessment Rubrics	Jamestown Journal - Group participation rating - Constructed response synthesis and interpretation.
Lesson 5 A Gift for Virginia	Text Materials Discussion Guide Editing Activity Guide Writing Assessment Rubrics	Jamestown Journal -Writing Activities Checklist -Editing Process Assessment -Trait based writing assessment

The Quality of the Jamestown Assessment Unit as Curriculum

The resources and extensive participation of curriculum experts and classroom teachers in the three year development cycle assured that the unit and lessons would have a user friendly teacher's guide and lessons that worked well with classroom groups of students. The quality was reflected in a number of formal and informal assessments of unit content.

- Reviews by Anchorage School District Reading, Language Arts, and Social Studies Curriculum Coordinators was positive.
- Reviews by Anchorage School District Curriculum Committees was positive and the unit and unit resources became part of the ASD approved supplemental curriculum materials.
- Reviews by Pilot Test teachers (14) were generally positive and, after minor revisions, the reviews by randomly selected classroom teachers (22) who taught the unit were positive.
- Review by scorers trained to use holistic and trait based rubrics for the independent scoring of student work indicated that students were responsive to the assignments given and that the work was scorable though minor revisions were required to the rubrics and work samples developed based on the trials with volunteer pilot group classes.

- Reviews by local university evaluators who are expert in Reading, Writing, and Social Studies instruction were positive.
- Review by the EXCELS! outside evaluator, Dr. Sandra Stotskey, was positive.

The Quality of the Jamestown Assessment Unit as Student Assessment

Assessment is the collection of information to be used in coming to understand and judge individuals, programs, and processes. The specific purposes of the Jamestown Unit were to report on the success of the EXCELS! project in improving the performance of students on skills called for in national, state, and local standards.

The validity of the assessment may be examined from three points of view. Because the assessment unit was accepted as part of the curriculum it is *ipso facto* an authentic assessment. The authenticity, however, goes beyond the content of the assessment to the activities that are entailed in the standard assessment pieces: listening and recording important facts, making a map which has acceptable form and includes the key features of Jamestown, reading multiple texts and demonstrating both specific recall of critical facts and the ability to make inferences and draw conclusions, performing well in a group working to solve an anthropological problem, developing a rational answer and justifying it, and writing a description and justification of an item which would transform and improve the lives of Jamestown residents of 1623. These activities are in and of themselves authentic classroom work and send the message to teachers and students that classroom work is important (See McTighe, 1996). Finally, the testimony of the various reviewers of the unit as quality instruction supports the link between the assessment activities and instruction.

An external indicator of the validity of the assessment is the relation of the scores on the Jamestown assessment to an independent measure of student performance on measures related to the same skills. Most of the students who participated in the Jamestown Assessment Unit classes were also tested with the California Achievement Test Survey Battery as part of the Alaska Statewide Student Assessment. Not all students were tested because of the normal turnover in students and Alaska Statewide Student Assessment regulations which allow the exclusion of some special education and bilingual students from testing.

The total number of students active in participating classes in the spring of 1997 was 412, including all bilingual and special education students. It was expected that the assessment would be given to bilingual and special education students with whatever accommodations were common in the classroom. One of the strengths of performance assessment units like

Jamestown is that it provides all students a chance to participate and can help support inclusion and equity in the classroom (Darling-Hammond, 1991).

The number of scorable responses on returned Jamestown Journals ranged from a low of 285 for the constructed response to reading following the New World News articles to 360 who answered at least one of the reading multiple choice items. Three hundred and sixty-eight of the 412 students, 89%, had at least one norm referenced test score. Table 1 shows the correlation of CAT 5 scores to the various included performance assessments based on the scoring of Jamestown Journals by trained raters.

The sampled classes were similar to the Anchorage School District as a whole. The average performances of the 368 students in the sample classes with test scores were compared to Anchorage School District averages on Reading, Total Language, Total Battery, and Social Studies the averages were not significantly different. The averages were, however, different among participating classrooms reflecting the diversity of the schools represented in Anchorage. Approximately 60% of the students where white, 10% were Alaska Native, 10% were Black and the remaining 20% were Pacific Islanders, Asian, and Hispanic. Forty-seven percent of the group were male. Twenty-seven percent qualified for free or reduced price lunches. Eight percent were participating in the special education program. Eight percent were participating in the bilingual program. Five students were getting help from migrant education program tutors. Class sizes ranged from 22 to 31 students with an average of 27.

All statistical analysis took place using the Statistical Package for the Social Sciences running on a VAX computer system. (SPSS, 1988). The data set is and a copy of the complete statistical analysis results are available on request from the authors.

When the correlations among the total scores for the various lessons are compared, they are also shown to be significantly associated though the degrees of association reflects notable differences. Lesson 2 (Listening and Map Making) scores correlate with lessons three, four and five .33*[2], .27*, .23*. Lesson 3 (Reading) scores correlate with lesson two, four, and five .33*, .17*, and .33*. Lesson 4 (Classification and Group Work) scores correlate with lessons two, three, and five .27*, .17* and .05. Lesson 4 (Writing) scores correlate with lessons two .23*, .33*, and .05. The individual lesson scores correlate with the overall total score .52*, .61*, .30*, and .85*.

[2] * is used throughout the paper to indicate that a result was found to be statistically significant with a probability level of .05 or lower.

Table 1
Correlation of Jamestown Journal and CAT 5 Scores

Jamestown Assessment	Average Score	Reading	Total Language	Total Battery	Social Studies
Total	2.66	.32*	.39*	.39*	.34*
Listening	2.48	.32*	.37*	.42*	.36*
Map	2.55	.28*	.32*	.38*	.23*
Reading -Multiple C.	3.04	.12*	.17*	.15*	.13*
Reading - Short Rsp	1.78	.15*	.29*	.29*	.31*
Catalogue Entry	3.56	.23*	.19*	.23*	.23*
Group Activity	2.68	.19*	.22*	.23*	.23*
Writing Ideas	2.87	.34*	.40*	.41*	.36*
Writing Conventions	2.85	.45*	.47*	.49*	.43*

While the correlations between the selected response norm referenced test and the performance assessment are not substantial, they are notable and statistically significant. While there was no specific expectation for agreement set prior to the development of the performance assessment and there was an expectation the differences would be notable, there was a hope that the correlations between the CAT Reading and Jamestown reading and CAT Total Language and the Jamestown writing assessment would be stronger than the correlations which were found.

The examination of the test-retest correlation of the CAT scores themselves were not much above .7 when the spring, 1996 and spring, 1997 CAT scores were examined for this group of students. The test-retest scores correlated .65* for Reading, .75* for Total Language, and .69* for Social Studies. The Total Battery test-retest correlation was .84*. About 300 of the 1997 students had scores for both 1996 and 1997.

Another question which arises with performance assessments is the consistency of scoring. The Jamestown Assessments were scored twice. The first scoring was done by the classroom teachers using the scales and score

point descriptions generated from the experience of the pilot trials. The second scoring was done by a group of trained raters who examined samples of the papers generated in 1997, selected samples of papers which they felt represented score points, modified the rubrics slightly based on the examination of the papers and discussion, and scored after demonstrating agreement in the rating of the sample exemplar papers. Rubrics were based on the Ideas and Content and Conventions rubrics developed by the Northwest Regional Educational Laboratory working with the Beaverton and Portland, OR school districts (Spandel and Culham, 1991). The trained raters were able to demonstrate at least 70% agreement in scoring papers.

Because of the process used to generate the expert scores, the agreement between teacher-raters and trained raters is examined as correlations rather than the number of absolute agreements.

Ratings by Teachers and Raters
Averages and Correlations

Area	Teacher Average	Trained Rater Average	Correlation
Total	34.59	33.09**	.86**
Listening	2.70	2.46**	.55**
Map	2.73	2,61*	.54**
Reading -Multiple C.	3.09	3.06	.95**
Reading - Short Rsp	2.33	1.84**	.57**
Catalogue Entry	4.12	3.58**	.27**
Group Activity	2.98	2.71**	.45**
Writing Ideas	2.86	2.90	.68*
Writing Conventions	2.89	2.86	.76*

***,** Represent significance less than or equal to .05 or .01 respectively. Means differences are based on t-test for paired samples and correlations are Pearson - r.**

All scoring guides are based on a 1 - 5 scale. The rubrics or scoring guides are adapted to the task and the social studies content of the artifact which is being scored. In general, a score of 1 represents a response which is "developing" or "beginning." A score of 1 is an active effort to respond to the assignment but may have little quality. A score of 3 is basic or a minimal demonstration that the standard has been fulfilled. A score of 3 represents an understanding of

the assignment and a performance which may have minor flaws and awkward constructions. A score of 5 represents advanced performance. Any flaws are so minor that they are easily passed over and the performance shows mastery of the content and skills covered in the assignment (Anchorage School District, 1995).

The average scores for the individual assessments are for the most part close to the score of 3. The short written response to reading which required the synthesis of ideas drawn from a series of articles was difficult. In most of the areas that did not require an assessment of writing, teachers gave higher scores than the trained raters. This was most notable on the catalogue entry which was a part of the "American Anthropologist" lesson. It may be that the completion of the lesson influenced the score for teachers while the trained raters only considered the quality of the entry and had the benefit of a set of exemplar papers to using. The correlation of catalogue entry scores was only .27.

However, the overall agreement of scores represented in the Total of individual assessments was .86 which reflects a substantial agreement between teachers and trained raters as to the overall quality of student performance. When this is contrasted with the correlation of the total score from the combination of the total score on the Jamestown performance assessment with the CAT/5 Total Battery score, .49*; Reading, .32*; Total Language, .40*; or Social Studies, .35*; it is fairly clear that what is being measured can be fairly consistently scored by teachers and trained raters while it is quite different from what is being measured by the selected response tests.

What does student performance on the Jamestown Assessment Unit mean for the EXCELS! program evaluation?

The goal is, of course to have all students performing at or above expectation which means having a score of 3, 4, or 5 on the various performance scales. There were two expectation set by the EXCELS! program: 1) that 70% of all students would meet or exceed district expectations and 2) student performance would improve. The table below examines the student performance in terms of program expectations.

The table shows the results from a trial of the 1995-1996 prototype Jamestown Assessment unit. Elements of the assessment were tested and scored on a trial basis for listening, 273 students; map making, 80 students; reading, 106 students, and writing, 50 students. The initial trial took place with volunteer teachers who were active participants in the EXCELS! project.

The 1996-1997 results are based on a random sample of intact grade 4 classroom groups. There was a potential of about 368 student participants

though not all students had provided scorable samples of work in all areas. The total included in each or the areas is listening, 314; map making, 328; reading, 328; group processes, 326; and writing 328. Students were some times excluded from parts of the assessment because they were special education or bilingual students but most of the missing scores were due to students who were absent on the day of a specific element. Average absences run between 5% and 8% of the student group. There was also one class excluded from the listening exercise because the teacher reported that she had not followed the directions provided with the lesson.

Percentage of Students Reaching Performance Goals
Jamestown Assessment Unit
1995-96 and 1996-97

Area	Selected Sample 1995-1996	Random Sample 1996-1997	Acceptable Performance
Listening	70%	49%	No
Map	85%	60%	No
Reading -Multiple Choice	82%	73%	Yes
Reading Written Response		24%	No
Catalogue Entry		86%	Yes
Writing Ideas	74%	64%	Yes/No
Writing Conventions	50%	61%	No

During these same two years, the norm referenced test scores in the areas of Reading, Total Language Arts, stayed the same or improved for fourth grade students. Reading scores were at the 64th percentile for both years. Total Language scores improved from the 56th to the 63rd percentile. The Total Battery score improved from the 63rd to the 66th percentile. Social Studies improved from the 57th to the 61st percentile.

When students with continuing enrollment with Anchorage School District were tested at the end of grade three and the end of grade 4 were compared there were also improvements. Reading improved from the 65th to the 66th

percentile. Total language arts improved from the 53rd to the 64th percentile. Social studies improved from the 58th to the 61st percentile. With over 2,800 students tested each year, all of the differences are statistically significant.

The Jamestown Assessment Unit and Cat Reading, Language Arts, and Social Studies scores suggest that students are making progress toward meeting Anchorage School District, state, and national goals for improved knowledge and skills. The Jamestown Assessment goes beyond the positive results of the standardized norm referenced tests to look at specific performance on skills and knowledge linked to the Anchorage curriculum. Taken together, the assessments are strong evidence of the success of the EXCELS! project.

Impact of the Assessment

The reaction of teachers to the Jamestown Assessment Unit was interesting. University observers, the programs's outside evaluator, and the Assessment and Evaluation team that created the Jamestown Assessment Unit visited classrooms and talked to teachers about the unit. There was also a brief survey with open-ended questions given to teachers along the assessment unit.

Nearly all of the teachers indicated that they enjoyed teaching Jamestown. They found that the materials were of high quality and that the assessments provided information that they could use on student performance. Some teachers extended the unit to include a lesson on map making after the map assessment and indicated that they were making modifications in their teaching to place more emphasis on listening and note taking. Some liked the editing process and editing check list included in the assignment and indicated that they would incorporate a similar peer editing and checklist activity into their teaching of writing. One teacher noted that it was the first time that he had curriculum materials which encouraged student research using the world wide web and found that his fourth grade students enjoyed using the Internet to gather information. It was evident that a number of teachers changed their instruction based on their experience with the Jamestown Assessment Unit.

It is clear that the Jamestown Assessment Unit in and of itself is a positive product of the EXCELS! project. It provides teachers with a model of a strong set of lessons based on an integration of social studies and language arts as well as effective student assessment. It also fulfills the goal of performance assessment that Khattri and all call "a seamless web that promotes teacher/student collaboration, active learning, critical thinking skills, and multi-disciplinary understanding." (Khattri, Kane and Reeve, 1995).

Conclusion

The production of the integrated Jamestown Assessment Unit was a successful part of EXCELS!. It demonstrates that student work generated as part of instruction may be used for a standardized assessment of student and program performance relative to local, state, and national standards.

The Jamestown Unit is an operational definition of an authentic assessment. The students learn and improve social studies and language arts skills through participation in the assessment unit. Teachers take an active part in the scoring of the assessments and have the option of grading their students relative to standards based expectations. Teachers also get feedback from trained scorers which allows them to consider their own grading and expectations for students. Curriculum developers and district administrators can see how students perform relative to model performances with known quality.

Analysis of the Jamestown Unit shows that performance assessment is scorable by both trained raters and teachers. The fairly similar scores and high overall correlation of total scores suggests that the unit may have achieved the goal of its developers to have a standardized assessment which could be fair as a basis for student grades.

While correlations with the Social Studies, Reading, and Language Arts sections of the CAT/5 survey battery are lower than expected by the authors, the correlations do show a strong significant relation. Based on reviews for validity of content and authenticity of performances by teachers, curriculum developers, and outside experts; our conclusion is that the Jamestown Assessment Unit measures somewhat different skills that are at least as well linked to standards as the norm referenced test questions.

Finally, a question of values.

The premise of the EXCELS! program was that integration of Language Arts and Social Studies instruction would enhance both curriculum areas. On the language arts side, students would become better communicators for their teachers would be self-conscious in providing experiences with reading, writing, speaking, and listening during social studies instruction. On the other social studies side, the integration of language arts activities into social studies content would give more time for social studies instruction and more emphasis on achieving the national, state, and local goals in civics, geography, history, and economics.

If you believe that teaching and learning is a self-conscious cycle of assessment, activity, and re-assessment of the knowledge and skills important to judging student success and adapting instruction; the Jamestown Unit may

have an importance beyond the EXCELS! project. It demonstrates that good curriculum and good assessment may be linked to curriculum through a set of authentic classroom instructional activities which provide for both learning and the creation of artifacts which may be used to judge student and program success. (See McTighe, 1996 and Wiggins, 1966).

The Jamestown Assessment Unit encourages good teaching and models good classroom assessment. It may be what we should expect to happen as we move forward into an era of standards driven instruction and assessment.

References

Alaska Department of Education. (June, 1996). Alaska Standards: Content Standards for Alaska Students. Juneau, AK: Alaska Department of Education.

Anchorage School District (1995) Jameston Integrated Assessment Unit Teacher's Guide. Anchorage, AK: Assessment and Evaluation Department, Anchorage School District.

Arter, J. (1991) Assessing Student Performance. Association for Supervision and Curriculum Development Inquiry Kit. Arlington, VA: The Association for Supervision and Curriculum Development.

ASCD. (1991). Designing Performance Assessment Tasks. Education Update Newsletter, 37 6, 1, 4-5, 8. Arlington, VA: The Association for Supervision and Curriculum Development.

Colonial National Historical Park (1995, August, 8). Classroom Enrichment at Jamestown and Yorktown. Yorktown, VA, Colonial National Historical Park.

Conover, K.A. (1996, November 19). Weekdays, She's in fifth Grade: Weekends, She's in 1627. The Christian Science Monitor, pp. 16, 17.

Falk, B and Darling-Hammond, L. (1993, March). The Primary Learning Record at P.S. 261: How Assessment Transforms Teaching and Learning." New York, NY: National Center for Restructuring Education, Schools and Teaching.

Fenton, R., Straugh, T., Ware, J, and. McPheters, S., and Stofflet, F. (1996). Anchorage School District Profile of Performance: 1997-98. Anchorage, AK: Anchorage School District, Assessment and Evaluation Department.

Fenton, R., Straugh, T., Garrison, S., Ware, J. McPheters, S., and Stofflet, F. (1997). Anchorage School District Profile of Performance: 1997-98. Anchorage, AK: Anchorage School District, Assessment and Evaluation Department.

James, S. (1996, November 19). A Visitor to Early Plymouth. The Christian Science Monitor, pp. 16, 17.

Jamestown-Yorktown Historical Society and Virginia Department of Education (Producer) ,Skinner, E. (Director). (Undated). Jamestown: The Beginning [Video]. (Available from Jamestown-Yorktown Foundation, PO Box 1607, Williamsburg, VA 23187-3630.

Khattri, N, Kane, M. and Reeve, A. (1995). Research Report: How Performance Assessments Affect Teaching and Learning. Educational Leadership, 53 3, 80-83.

Koretz, D., Stecher, B., Klein, D., McCaffrey, D. and Diebert, D. (1993, December). Can Portfolios Assess Student Performance and Influence Instruction? The 1991-1992 Vermont Experience. CSE Technical Report 371. Los Angeles, CA: CRESST.

McTighe, J. (1966, December). What Happens Between Assessments? Educational Leadership, 54, 6-12.

National Park Service (Undated). Jamestown, Where a Nation Began. Richmond, VA, Association for the Preservation of Virginia Antiquities.

Spandel, V and Culham, R. (1991). The Student Friendly Guide to Writing with Traits. Portland, OR: Northwest Regional Educational Laboratory.

Pearson, J. and Watkins, B. (1991). Early Jamestown: A Unit of Study for Grades 5-8. Los Angeles, CA, National Center for History in the Schools, University of California, Los Angeles.

Shavelson, R. (1995, June 2). Notebook: Development of Performance Assessments in Science - A Workshop Designed for the Earth Systems Implementation Program, Anchorage School District. Palo Alto, CA, Stanford University.

Sparks, M. (1996, November 19). Unravel These 17th-Century Riddles from the Book of Merry Riddles. The Christian Science Monitor, pp. 16, 17.

SPSS Inc, (1988). SPSS-X User's Guide. Chicago, IL, SPSS Inc.

Thomas, O. (1996, November 19). What did Pilgrim Children Do Without Videos. The Christian Science Monitor, pp. 16, 17.

Wiggins, G. (1966). Practicing What We Preach in Designing Authentic Assessments, Educational Leadership, 54, 18-25.